



RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/759,334A
Source: IFWO
Date Processed by STIC: 6/15/04

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER
VERSION 4.2 PROGRAM ACCESSIBLE THROUGH THE U.S. PATENT AND
TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker/chkrnote.htm>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry, Federal Express, United Parcel Service, or other delivery service (EFFECTIVE 06/05/04):
U.S. Patent and Trademark Office, 220 20th Street S., Customer Window, Mail Stop Sequence, Crystal Plaza Two, Lobby,
Room 1B03, Arlington, VA 22202

Raw Sequence Listing Error Summary

ERROR DETECTED

SUGGESTED CORRECTION

SERIAL NUMBER: 10759,334A

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 Wrapped Nucleics
 Wrapped Aminos The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2 Invalid Line Length The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3 Misaligned Amino
 Numbering The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4 Non-ASCII The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5 Variable Length Sequence(s) contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6 PatentIn 2.0
 "bug" A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7 Skipped Sequences
 (OLD RULES) Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence
 (2) INFORMATION FOR SEQ ID NO: X: (insert SEQ ID NO where "X" is shown)
 (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading).
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: X: (insert SEQ ID NO where "X" is shown)
 This sequence is intentionally skipped

 Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8 Skipped Sequences
 (NEW RULES) Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence
 <210> sequence id number
 <400> sequence id number
 000
- 9 Use of n's or Xaa's
 (NEW RULES) Use of n's and/or Xaa's have been detected in the Sequence Listing.
 Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
 In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 10 Invalid <213>
 Response Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
- 11 ✓ Use of <220> Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown". Please explain source of genetic material in <220> to <223> section.
 (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12 PatentIn 2.0
 "bug" Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13 Misuse of n/Xaa "n" can only represent a single nucleotide; "Xaa" can only represent a single amino acid



IFWO

RAW SEQUENCE LISTING

DATE: 06/15/2004

PATENT APPLICATION: US/10/759,334A

TIME: 17:09:49

Input Set : A:\Pruitt Cell Lineage sequence listing Non-Prov.txt

Output Set: N:\CRF4\06152004\J759334A.raw

3 <110> APPLICANT: Pruitt, Steven C.

5 <120> TITLE OF INVENTION: Method for Comprehensive Identification of Cell Lineage
Specific Genes

7 <130> FILE REFERENCE: 03551.0124

9 <140> CURRENT APPLICATION NUMBER: 10/759,334A

10 <141> CURRENT FILING DATE: 2004-01-16

12 <150> PRIOR APPLICATION NUMBER: US 60/440,510

13 <151> PRIOR FILING DATE: 2003-01-16

15 <160> NUMBER OF SEQ ID NOS: 39

17 <210> SEQ ID NO: 1

18 <211> LENGTH: 151

19 <212> TYPE: DNA

20 <213> ORGANISM: artificial sequence

22 <220> FEATURE:

23 <221> NAME/KEY: unsure

24 <222> LOCATION: 19-32, 51-64, 83-96, 103-116, 147-151

25 <223> OTHER INFORMATION: n is a,t,g or c

27 <400> SEQUENCE: 1

-> 29 tctagacagt ctggagagnn nnnnnnnnnn nntctagaca gtctggagag 50.

-> 30 nnnnnnnnnn nnnntctaga cagtctggag agnnnnnnnn nnnnnntcta 100

-> 31 gannnnnnnn nnnnnnctct ccagactgtc tagacagtct ggagagnnnn 150

-> 32 n 151

34 <210> SEQ ID NO: 2

35 <211> LENGTH: 18

36 <212> TYPE: DNA

37 <213> ORGANISM: artificial sequence

39 <220> FEATURE:

40 <223> OTHER INFORMATION: PCR Primer

42 <400> SEQUENCE: 2

43 ccgcttttct ggattcat 18

45 <210> SEQ ID NO: 3

46 <211> LENGTH: 21

47 <212> TYPE: DNA

48 <213> ORGANISM: artificial sequence

50 <220> FEATURE:

51 <223> OTHER INFORMATION: MAGE Universal Adapter

53 <400> SEQUENCE: 3

54 tctagaggac tgcgtggcg a 21

56 <210> SEQ ID NO: 4

57 <211> LENGTH: 25

58 <212> TYPE: DNA

59 <213> ORGANISM: artificial sequence

61 <220> FEATURE:

62 <221> NAME/KEY: unsure

Does Not Comply
Corrected Diskette Needed
(pg. 1)please
explainplease see item
11 on error
summary sheet.

RAW SEQUENCE LISTING

DATE: 06/15/2004

PATENT APPLICATION: US/10/759,334A

TIME: 17:09:49

Input Set : A:\Pruitt Cell Lineage sequence listing Non-Prov.txt

Output Set: N:\CRF4\06152004\J759334A.raw

63 <222> LOCATION: 24,25
64 <223> OTHER INFORMATION: MAGE Universal Adapter; n is a,t,g or c
66 <400> SEQUENCE: 4
67 cctcgcccccac gcagtcctct agann 25
69 <210> SEQ ID NO: 5
70 <211> LENGTH: 19
71 <212> TYPE: DNA
72 <213> ORGANISM: artificial sequence
74 <220> FEATURE:
75 <223> OTHER INFORMATION: MAGE PCR Primer
77 <400> SEQUENCE: 5
78 cctcgcccccac gcagtcctc 19
80 <210> SEQ ID NO: 6
81 <211> LENGTH: 18
82 <212> TYPE: DNA
83 <213> ORGANISM: artificial sequence
85 <220> FEATURE:
86 <223> OTHER INFORMATION: MAGE PCR Primer
88 <400> SEQUENCE: 6
89 cggctgggtg tggcggac 18
91 <210> SEQ ID NO: 7
92 <211> LENGTH: 16
93 <212> TYPE: DNA
94 <213> ORGANISM: Mus musculus
96 <400> SEQUENCE: 7
97 ggcgacacgc gcacct 16
99 <210> SEQ ID NO: 8
100 <211> LENGTH: 16
101 <212> TYPE: DNA
102 <213> ORGANISM: Mus musculus
104 <400> SEQUENCE: 8
105 agggaggaga tgtagt 16
107 <210> SEQ ID NO: 9
108 <211> LENGTH: 16
109 <212> TYPE: DNA
110 <213> ORGANISM: Mus musculus
112 <400> SEQUENCE: 9
113 actacatctc ctccct 16
115 <210> SEQ ID NO: 10
116 <211> LENGTH: 16
117 <212> TYPE: DNA
118 <213> ORGANISM: Mus musculus
120 <400> SEQUENCE: 10
121 gggttgtctt cactct 16
123 <210> SEQ ID NO: 11
124 <211> LENGTH: 16
125 <212> TYPE: DNA
126 <213> ORGANISM: Mus musculus
128 <400> SEQUENCE: 11

RAW SEQUENCE LISTING

DATE: 06/15/2004

PATENT APPLICATION: US/10/759,334A

TIME: 17:09:49

Input Set : A:\Pruitt Cell Lineage sequence listing Non-Prov.txt

Output Set: N:\CRF4\06152004\J759334A.raw

```
129 cccaggatgt atagct 16
131 <210> SEQ ID NO: 12
132 <211> LENGTH: 16
133 <212> TYPE: DNA
134 <213> ORGANISM: Mus musculus
136 <400> SEQUENCE: 12
137 agtctggaaa cctgtc 16
139 <210> SEQ ID NO: 13
140 <211> LENGTH: 16
141 <212> TYPE: DNA
142 <213> ORGANISM: Mus musculus
144 <400> SEQUENCE: 13
145 actacatctc ctacct 16
147 <210> SEQ ID NO: 14
148 <211> LENGTH: 16
149 <212> TYPE: DNA
150 <213> ORGANISM: Mus musculus
152 <400> SEQUENCE: 14
153 aggtaagtgc tgctgc 16
155 <210> SEQ ID NO: 15
156 <211> LENGTH: 16
157 <212> TYPE: DNA
158 <213> ORGANISM: Mus musculus
160 <400> SEQUENCE: 15
161 aggtaagtac aagctg 16
163 <210> SEQ ID NO: 16
164 <211> LENGTH: 16
165 <212> TYPE: DNA
166 <213> ORGANISM: Mus musculus
168 <400> SEQUENCE: 16
169 acctcagttg attcct 16
171 <210> SEQ ID NO: 17
172 <211> LENGTH: 16
173 <212> TYPE: DNA
174 <213> ORGANISM: Mus musculus
176 <400> SEQUENCE: 17
177 agctttgaat tcatga 16
179 <210> SEQ ID NO: 18
180 <211> LENGTH: 16
181 <212> TYPE: DNA
182 <213> ORGANISM: Mus musculus
184 <400> SEQUENCE: 18
185 tttctcaggg tagcct 16
187 <210> SEQ ID NO: 19
188 <211> LENGTH: 16
189 <212> TYPE: DNA
190 <213> ORGANISM: Mus musculus
192 <400> SEQUENCE: 19
193 tccgctcaat gtacct 16
```

RAW SEQUENCE LISTING

DATE: 06/15/2004

PATENT APPLICATION: US/10/759,334A

TIME: 17:09:49

Input Set : A:\Pruitt Cell Lineage sequence listing Non-Prov.txt

Output Set: N:\CRF4\06152004\J759334A.raw

195 <210> SEQ ID NO: 20
196 <211> LENGTH: 16
197 <212> TYPE: DNA
198 <213> ORGANISM: Mus musculus
200 <400> SEQUENCE: 20
201 agggttggac tcaagg 16
203 <210> SEQ ID NO: 21
204 <211> LENGTH: 16
205 <212> TYPE: DNA
206 <213> ORGANISM: Mus musculus
208 <400> SEQUENCE: 21
209 actacatctc ctccgt 16
211 <210> SEQ ID NO: 22
212 <211> LENGTH: 16
213 <212> TYPE: DNA
214 <213> ORGANISM: Mus musculus
216 <400> SEQUENCE: 22
217 cacggggggcg gagcct 16
219 <210> SEQ ID NO: 23
220 <211> LENGTH: 16
221 <212> TYPE: DNA
222 <213> ORGANISM: Mus musculus
224 <400> SEQUENCE: 23
225 aggtaaccct gtgctg 16
227 <210> SEQ ID NO: 24
228 <211> LENGTH: 16
229 <212> TYPE: DNA
230 <213> ORGANISM: Mus musculus
232 <400> SEQUENCE: 24
233 aggtagacta cttgtg 16
235 <210> SEQ ID NO: 25
236 <211> LENGTH: 16
237 <212> TYPE: DNA
238 <213> ORGANISM: Mus musculus
240 <400> SEQUENCE: 25
241 cgccatcgct ctagct 16
243 <210> SEQ ID NO: 26
244 <211> LENGTH: 16
245 <212> TYPE: DNA
246 <213> ORGANISM: Mus musculus
248 <400> SEQUENCE: 26
249 aggctcatca gctgac 16
251 <210> SEQ ID NO: 27
252 <211> LENGTH: 16
253 <212> TYPE: DNA
254 <213> ORGANISM: Mus musculus
256 <400> SEQUENCE: 27
257 aggtgtggat ccagag 16
259 <210> SEQ ID NO: 28

RAW SEQUENCE LISTING

DATE: 06/15/2004

PATENT APPLICATION: US/10/759,334A

TIME: 17:09:49

Input Set : A:\Pruitt Cell Lineage sequence listing Non-Prov.txt

Output Set: N:\CRF4\06152004\J759334A.raw

260 <211> LENGTH: 16
261 <212> TYPE: DNA
262 <213> ORGANISM: Mus musculus
264 <400> SEQUENCE: 28
265 aggtaggtct cgatct 16
267 <210> SEQ ID NO: 29
268 <211> LENGTH: 16
269 <212> TYPE: DNA
270 <213> ORGANISM: Mus musculus
272 <400> SEQUENCE: 29
273 agaatacgat acccag 16
275 <210> SEQ ID NO: 30
276 <211> LENGTH: 16
277 <212> TYPE: DNA
278 <213> ORGANISM: Mus musculus
280 <400> SEQUENCE: 30
281 agggcatttc tattac 16
283 <210> SEQ ID NO: 31
284 <211> LENGTH: 16
285 <212> TYPE: DNA
286 <213> ORGANISM: Mus musculus
288 <400> SEQUENCE: 31
289 agctaatacac caaagc 16
291 <210> SEQ ID NO: 32
292 <211> LENGTH: 16
293 <212> TYPE: DNA
294 <213> ORGANISM: Mus musculus
296 <400> SEQUENCE: 32
297 aggatccact gtttac 16
299 <210> SEQ ID NO: 33
300 <211> LENGTH: 16
301 <212> TYPE: DNA
302 <213> ORGANISM: Mus musculus
304 <400> SEQUENCE: 33
305 aggtcctgcc tcagca 16
307 <210> SEQ ID NO: 34
308 <211> LENGTH: 16
309 <212> TYPE: DNA
310 <213> ORGANISM: Mus musculus
312 <400> SEQUENCE: 34
313 aggtcctgcc tcagca 16
315 <210> SEQ ID NO: 35
316 <211> LENGTH: 16
317 <212> TYPE: DNA
318 <213> ORGANISM: Mus musculus
320 <400> SEQUENCE: 35
321 aggactgatt gtggtg 16
323 <210> SEQ ID NO: 36
324 <211> LENGTH: 16

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/759,334A

DATE: 06/15/2004
TIME: 17:09:50

Input Set : A:\Pruitt Cell Lineage sequence listing Non-Prov.txt
Output Set: N:\CRF4\06152004\J759334A.raw

Note:

n and/or Xaa have been detected in the Sequence Listing. Please review the
e Listing to ensure that a corresponding explanation is presented in the <220>
> fields of each sequence which presents at least one n or Xaa.

N Pos. 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 51, 52, 53, 54, 55, 56
N Pos. 57, 58, 59, 60, 61, 62, 63, 64, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94
N Pos. 95, 96, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115
N Pos. 116, 147, 148, 149, 150, 151
N Pos. 24, 25

VERIFICATION SUMMARY

DATE: 06/15/2004

PATENT APPLICATION: US/10/759,334A

TIME: 17:09:50

Input Set : A:\Pruitt Cell Lineage sequence listing Non-Prov.txt

Output Set: N:\CRF4\06152004\J759334A.raw

L:29 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:0
L:30 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:50
L:31 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:100
L:32 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:150
L:67 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:0